DOCUMENT RESUME

ED 285 378 EC 200 467

AUTHOR Martindale, Elizabeth S.; And Others
TITLE Accuracy of Class.LD2: An Expert System for Classifying Learning Disabled Students.
PUB DATE 86
NOTE 34p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Classification; Computer Managed Instruction; *Computer Uses in Education; Decision Making; Elementary Secondary Education; *Eligibility; *Expert Systems; *Learning Disabilities; *Student Placement
IDENTIFIERS *Student Placement Teams

ABSTRACT

The study evaluated the accuracy of Class.LD2, a computerized expert system for classifying learning disabled (LD) students. Of 264 student files, 78 files were chosen based on disagreement between multidisciplinary team and Class.LD2 decisions regarding eligibility for LD placement. These 78 cases were evaluated by three individuals expert in LD classification, who made an eligibiltiy decision for each case. Analysis of the results indicated that: (1) Class.LD2 was in agreement with the experts more often than were the placement teams; (2) the expert system's decisions were significantly correlated with those of the experts; and (3) in those cases in which the experts were in unanimous disagreement with Class.LD2, it was shown that Class.LD2 conformed strictly to state rules and regulations in making its eligibility decisions. The use of Class.LD2 by placement teams in order to encourage more data based decisions and to limit overclassification errors is also discussed. (Author/DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED285378

Accuracy of Class.LD2: An Expert System for Classifying Learning Disabled Students

Elizabeth S. Martindale, Joseph M. Ferrara, and Brian W. Campbell

Utah State University

Department of Special Education

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Elizabeth Martindale

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Running head: ACCURACY OF CLASS.LD2

EC200467

Accepted for publication in Computers in Human Behavior (in press)

Abstract

This study determined the accuracy of Class.LD2, an expert system for classifying learning disabled students. Of 264 student files, 78 files were chosen based on disagreement between multidisciplinary team and Class.LD2 decisions regarding eligibility for LD placement. These 78 cases were evaluated by three individuals expert in LD classification, who made an eligibility decision for each case. Their decisions were compared to those of the expert system2. Analysis of the results indicated that 1) Class.LD2 was in agreement with the experts more often than were the placement teams, 2) the expert system's decisions were significantly correlated with those of the experts, and 3) in those cases in which the experts were in unanimous disagreement with Class.LD2, it was shown that Class.LD2 conformed strictly to state rules and regulations in making its eligibility decisions. The use of Class.LD2 by placement teams in order to encourage more data based decisions and to limit overclassification errors is also discussed.

Accuracy of Class.LD2: an Expert System for Classifying Learning Disabled Students

Special education placement teams must make decisions regarding the identification of handicapped children and their eligibility for special education services. Current data suggests the judgments made by placement teams may not be accurate. Ysseldyke (1983) contends that half of the number of identifications in the area of learning disabilities may be inaccurate. Hofmeister (1983) reported an 84% increase in the number of pupils identified as learning disabled during the past few years. Placement team inaccuracy may be related to this dramatic increase in the number of pupils identified as learning disabled (Algozzine & Korinek, 1985; Algozzine, Ysseldyke & Christenson, 1983).

If placement teams are doing a poor job of applying eligibility criteria, handicapped and nonhandicapped children are not being well served. First, nonhandicapped children may be receiving inappropriate services. Second, if nonhandicapped students are qualifying for special education, there is less money available to meet the needs of the

handicapped, and essential services for children with other handicaps are being reduced to pay for the new learning disabled students (Sabatino, 1981).

In some states the problem of overclassification has reached crisis proportions. State education agencies have responded to this problem by placing limits on the number of students who can qualify and receive funding for services in the learning disabled category (Boyan, 1985; Reynolds, 1983). Such limits may solve the immediate fiscal crisis. Unfortunately, they subvert the intent of P.L. 94-142 in that they discriminate against handicapped children who happen to live in schools with an unusually high handicapped population. At the same time, they encourage aggressive and "imaginative" administrators in districts with a low percentage of handicapped students to find pseudo-learning disabled students to fill their "quota" and obtain additional funding.

An alternative to state imposed limits is to improve the accuracy of the decisions made by placement teams. Research on the current functioning of placement teams indicates that their inaccuracy may be the result of a flawed approach to placement

decision making. Few school placement teams employ a systematic approach to determining eligibility (Ysseldyke, Algozzine & Mitchell, 1982). As a result, a great deal of data describing student performance is collected, but much of it is technically irrelevant (Thurlow & Ysseldyke, 1979). Teams spend about 30% of their time discussing these questionable data (Ysseldyke, Algozzine, Rostollan, & Shinn, 1981). Furthermore, individuals discussing these data are likely to use language which is unfamiliar to noneducators (Ysseldyke, 1983). Considering this unfortunate state of affairs, it is not surprising that Ysseldyke, Algozzine, Richey, and Graden (1982) reported that there was little relationship between the psychometric data presented to placement teams and the eligibility decisions made by those teams.

Two changes in the decision making process might reduce the frequency of errors. First, teams should follow a rational series of steps which encourage the examination of relevant data. Second, teams should use a set of regulation-based rules to make judgments about placement data.

One way to bring about these changes involves the application of expert system technology. An expert system is a computer program which attempts to replicate the decision making and problem solving skills of knowledgeable and effective human experts (Hofmeister & Ferrara, in press). If an expert system which behaved like a knowledgeable, systematic expert in the area of eligibility decision making could be developed, it might serve as a useful tool for placement teams. An expert system might force a rational step-by-step approach which would result in judgments which were directly related to data and regulations (Hofmeister & Lubke, in press).

Before such a system can be used by placement teams its judgement must be shown to be competent. An expert system, Class.LD2 (Ferrara & Hofmeister, 1984), has been developed at Utah State University. The purpose of Class.LD2 is to evaluate student eligibility for special education placement (Ferrara, Parry & Lubke, 1985). The degree to which Class.LD2 outperforms placement teams should be related to its potential value as a tool for increasing the accuracy of placement team decisions. One way to evaluate

Class.LD2's performance is to compare its judgments to those of placement teams as well as to the judgments of acknowledged experts in the area of LD classification and placement. The purpose of this study was to make these comparisons.

This study addressed the following research questions:

1. Were the decisions made by Class.LD2 and the placement teams different?

2. When placement teams and Class.LD2 did not agree, what was the nature of the disagreement?

3. In cases where placement teams and Class.LD2 disagreed, whose decisions did experts judge to be correct?

4. What were the characteristics of cases where experts agreed with the placement team instead of Class.LD2?

METHOD

Class.LD2

Class.LD2 (Ferrara & Hofmeister, 1984) is an expert system designed to provide a second opinion regarding the classification of students as learning disabled (Ferrara et al., 1985; Hofmeister &

Lubke, in press). Class.LD2 uses a dialogue or consultation format to obtain information from the user. That information is compared against the rules in a preprogrammed knowledge base and is used to make a classification decision.

CLASS.LD2's knowledge base contains approximately 600 "if-then" rules. The rules are based on P.L. 94-142, Utah state regulations, and current literature describing best practices in the area of student assessment and placement. As questions are generated by the program and answered by the user, decisions are made based on a comparison of the answers to the information in the knowledge base.

Sample cases

Data from 264 files which described potential special education students were used to evaluate the performance of Class.LD2. These files were selected from three school districts. District 1, located in Idaho, provided 150 files. District 2, located in Utah, contributes 54 files, and District 3, located in Wyoming, contributed 60 files. Of the total number of files, 110 of the students were judged by placement teams to be eligible for placement as learning

disabled. The remaining 154 cases were either judged to be nonhandicapped or eligible for services in another disability area.

Data Input

The data from each student file was provided to Class.LD2 in a consultation format. Special education teachers familiar with each case entered the data from Districts 2 and 3. District 1 data were entered by graduate students pursuing a PhD in psychology. The graduate students worked as consultants to District 1 and were familiar with each case. Neither the graduate students nor the teachers were involved in the development of Class.LD2. They were provided instruction on Class.LD2 prior to their use of the system. Records of each data entry session were kept. These records contained the questions asked by Class.LD2 and each response made by the users. Data entry sessions took about 8 minutes per case.

Disagreeing Cases

When Class.LD2 had evaluated each student's data, its judgment was compared with the judgment made by the placement team. Cases where disagreement occurred were identified.

Class.LD2 does not normally provide a dichotomous judgment (i.e., LD or not LD). Rather, Class.LD2 will provide as many as eight positive hypotheses regarding the student's area of disability. Certainty factors of -100 through +100 are then assigned to these hypotheses. In order to judge each case as an agreement or disagreement, Class.LD2's judgments had to be viewed as dichotomous rather than continuous data. If Class.LD2 assigned a certainty factor of +50 or greater to the LD hypothesis, the student was viewed as being judged by the system to be eligible for LD services. Therefore, if a certainty level of +49 or lower was assigned, the student was considered to be not LD.

Expert Evaluator

Cases where Class.LD2 and the placement team disagreed were submitted to three experts. Expert 2 is a school psychologist in a large Utah school district. Expert 1 and Expert 3 are employed by the Utah State Office of Education. A principle component of all three experts' work involves assisting local school districts in making placement decisions.

The experts were given data sheets which

contained the same information that was provided to Class.LD2. Based on this information, the experts were asked to judge students as either eligible or ineligible for LD placement.

In cases where the experts unanimously disagreed with Class.LD2's judgement, the intermediate conclusions of Class.LD2 were examined. Specifically, three critical attributes were examined: (a) IQ level, (b) a discrepancy score based on the student's actual grade level and, (c) a discrepancy score based on an estimate of expected performance.

Results

Placement Team Decisions

Of the 264 files checked with Class.LD2, there were 78 cases where placement team decisions did not match Class.LD2's decisions. In other words, Class.LD2 and the placement teams were in agreement 67% of the time. Table 1 shows a 2x2 table illustrating the relationship between Class.LD2's decisions and the placement teams' decisions. A phi coefficient of phi=.46 was calculated using these data.

Insert Table 1. about here

Tables 2, 3, and 4 show the relationship between the placement decisions of Districts 1, 2, and 3 and those of Class.LD2. A phi coefficient of phi=.27 was calculated for District 1. A phi coefficient of phi=.56 was obtained for District 2. Finally, the relationship between Class.LD2's judgments and those of District 3's placement team is described by a phi=.24.

Insert Table 2. about here

Insert Table 3. about here

Insert Table 4. about here

Nature of Disagreements

There were 78 cases where Class.LD2 and the placement team did not agree. These 78 cases can be divided into two categories: (a) Type A cases, where the placement team said the student was learning disabled and Class.LD2 said the student was not, and (b) Type B cases, where the placement team said the student was not learning disabled and Class.LD2 said

the student was learning disabled. There were 68 Type A cases and 10 Type B Cases. When disagreement occurred, then, 87% of the disagreements were a result of either team overclassification or machine underclassification.

Accuracy of Decisions

The 78 cases where disagreement occurred were submitted to three experts. When taken together, the experts made 234 (3X78) judgments. Figure 1 shows the experts' total agreement and disagreement with Class.LD2.

Insert Figure 1. about here

Table 5 shows the relationship between the experts' placement decisions and those of Class.LD2. Using the data shown in Table 5, a phi coefficient of phi=.40 was calculated. This phi is associated with a chi-square of 36.62, which is significant at the level required to reject the null hypothesis that there is not a relationship between the experts' judgments and Class.LD2's at the $p<.001$ level of confidence.

Insert Table 5. about here

Table 6 shows the relationship between Class.LD2's decisions and all cases where the three experts agreed. There were 58 such cases. Phi=.63 and a chi-square of 23.38 were calculated to describe this relationship. Once again we can reject the null hypothesis that there was no relationship between Class.LD2's judgments and the experts' judgments at the $p<.001$ level of confidence.

Insert Table 6. about here

Table 7 shows a correlation matrix describing the relationship between all five judgments of student eligibility (Class.LD2, placement teams, and the experts). The decisions of Class.LD2 and of all three experts are inversely related to placement team decisions. It should be noted that only cases where placement teams and Class.LD2 disagreed were used to calculate these correlations.

Insert Table 7. about here

Characteristics of Cases

The human experts' judgments were unanimous in only 6 of the 78 cases of Class.LD2/placement team disagreements. Table 8 shows Class.LD2's intermediate conclusions for these cases on their critical variables: IQ, estimated discrepancy and actual discrepancy. In all six of the cases where Class.LD2 and the experts disagreed, Class.LD2 made an underclassification error.

Insert Table 8. about here

In each case, the student's data failed to satisfy one of the LD cutoff levels established in Class.LD2's knowledge base. These cutoff levels include a minimum IQ of 83 or 84 (depending on the standard deviation of the IQ test used) and a 40% discrepancy between expected and observed academic performance.

Discussion

Overall Agreement

The results of this study suggest that Class.LD2 and placement teams tended to agree on most (74%) cases. There was, however, enough disagreement to

justify an evaluation of the 78 disagreeing cases.

Evaluation of Disagreeing Cases

The 78 cases in which Class.LD2 and the placement teams disagreed can be viewed as difficult cases. In these situations Class.LD2 clearly and consistently outperformed the placement teams. The few cases in which the experts all disagreed with Class.LD2 were marginal. Furthermore, Class.LD2's decisions were defensible and its reasons for making those decisions were clear.

Cautions for Interpreting the Data

This study uses a portion of the data which were collected during the formative evaluation of Class.LD2. As a result, both the data and the analysis are geared toward product development rather than a summative report on the performance of the product. The reader should be alerted to a number of conditions which limit the degree to which the conclusions of this study can be assumed to generalize to a larger population.

Utah Guidelines for Idaho and Wyoming Students

The reader should recognize that Utah's LD classification criteria were used to evaluate files

from two other states. Class.LD2 was more likely to agree with placement teams from District 2 (the Utah District) than with teams from the Idaho and Wyoming school districts. The placement teams' degree of inaccuracy which might be inferred from these data is probably inflated.

Fence Sitting was not Allowed

In this analysis, Class.LD2's advice was transformed from continuous data to dichotomous data. As a result, some error was introduced. For example, Class.LD2's consideration of one student's data resulted in an advice statement which suggested that the student could be classified LD at a +48 level of confidence. Using a +50 cutoff, Class.LD2's advice in that case was considered to be not LD. In the field, a +48 certainty level would probably not be interpreted as an absolutely negative response. Most professionals would recognize that Class.LD2 was suggesting that, in this case, an LD placement would be marginally defensible. The continuous nature of Class.LD2's advice would make its judgments useful even in those cases where it disagreed with one or more of the experts. But in this study near misses

did not count.

Controversial Issues: Class.LD2 is Smarter than it Looks

Several elements of the currently accepted criteria for LD placement are open to professional debate. For example, if a student's learning problems can be primarily attributed to environmental, economic, or cultural deprivation, that student cannot be classified as learning disabled.

Not everyone agrees (Sabatino, 1983) that students with learning problems attributable to environmental or economic factors should be excluded from LD placement. Conversely, concern has also been expressed that environmentally and economically disadvantaged students have been placed in special education programs and that this practice could place learning disability programs in danger of becoming a dumping ground for all educational problems (Kirk & Kirk, 1983).

The experts evaluating the placement team/Class.LD2 disagreement cases were not in agreement on the environmental or economic issue. In nine cases where there was clear environmental and/or economic deprivation, the experts disagreed. Two of

the experts chose to call all nine of these children learning disabled. The other expert chose not to classify any of them as learning disabled.

In all of these cases, Class.LD2's advice took a middle ground. The program alerted its users to the controversial issue and then pointed out that a strict interpretation of current state and federal guidelines would not allow an LD placement. For the purpose of analysis, Class.LD2's advice in all such cases was judged as not LD. This interpretation accounted for a large portion of the disagreements between Class.LD2 and the experts. Clearly, the analysis conducted in this study was not sensitive to Class.LD2's appropriately moderate advice in controversial areas.

Use of An Expert System to Limit Overclassification

If Class.LD2 had been used by placement teams to generate a second opinion regarding LD placement, the number of overclassification errors might have been reduced. The decisions made by placement teams might then have reflected a more accurate application of state and federal rules and regulations.

If the use of Class.LD2 does indeed reduce placement team error, the need for state imposed LD

limits will be reduced. Current research is attempting to measure the effect of Class.LD2 on placement team performance.

References

Algozzine, B. & Korinek, L. (1985). Where is special education going? Exceptional Children, 51 (5), 388-394.

Algozzine, B., Ysseldyke, J. E., & Christenson, S. (1983). An analysis of the incidence of special class placement: The Masses are burgeoning. The Journal of Special Education, 17 (2), 141-147.

Boyan, C. (1985). California's new eligibility criteria: Legal and program implications. Exceptional Children, 52 (2), 131-141.

Ferrara, J. M., & Hofmeister, A. M. (1984). CLASS.LD2: An expert system for classifying learning disabilities, [Computer program]. Logan, UT: Utah State University.

Ferrara, J. M., Parry, J. D., & Lubke, M. M. (1985). CLASS.2: An expert system for student classification. Technical Paper. Logan, UT: Utah State University. (ERIC Document Reproduction Service No. ED 263 734)

Hofmeister, A. M. (March 21, 1983). Comments on part B, Section 618: Evaluation and Part E, research and education of the handicapped, Public Law 94-142.

Testimony before the Senate Committee on Labor and Human Resources, Subcommittees on the Handicapped, Congressional Record. Washington, DC: Government Printing Office.

Hofmeister, A. M., & Ferrara, J. M. (in press). Expert systems and special education. Exceptional Children.

Hofmeister, A. M., & Lubke, M. M. (in press). Expert systems: Implications for the diagnosis and treatment of learning disabilities. Learning Disabilities Quarterly.

Kirk, S. A., & Kirk, W. D. (1983). On defining learning disabilities. Annual Review of Learning Disabilities 1, 17-18.

Reynolds, M. C. (1983). Classification of students with handicaps. Unpublished manuscript, University of Minnesota, Minneapolis.

Sabatino, D. A. (1981). Are appropriate educational programs operationally achievable under mandated promises of P.L. 94-142? The Journal of Special Education, 15, 9-22.

Sabatino, D. A. (1983). The house that Jack built. Annual Review of Learning Disabilities 1, 23-24.

Thurlow, M., & Ysseldyke, J. E. (1979). Current assessment and decision-making practices in model programs for learning disabled students. Learning Disability Quarterly, 2, 15-24.

Ysseldyke, J. E. (1983). Current practices in making psychoeducational decisions about learning disabled students. Annual Review of Learning Disabilities, 1, 31-38.

Ysseldyke, J. E., Algozzine, B., & Mitchell, J. (1982). Regular education teacher participation in special education team decision making. Elementary School Journal, 60, 308-313.

Ysseldyke, J. E., & Algozzine, B., Richey, L., & Graden, J. (1982). Declaring students eligible for learning disability services: Why bother with the data? Learning Disability Quarterly, 5, 37-44.

Ysseldyke, J. E., Algozzine, B., Rostollan, D., & Shinn, M. (1981). A content analysis of the data presented at special education team meetings. Journal of Clinical Psychology, 37, 665-662.

Table 1

Class.LD2 and overall Placement Team Decisions

CLASS.LD2	PLACEMENT TEAMS: LD	PLACEMENT TEAMS: NOT LD
LD	64	10
NOT LD	68	122

N = 264

Total Agreement = 64 + 122 = 186

Total Disagreement = 10 + 68 = 78

Phi Coefficient = .46

Table 2

Class.LD2 and District 1 Placement Team Decisions

		PLACEMENT TEAMS: LD	PLACEMENT TEAMS: NOT LD
CLASS.LD2	LD	32	6
	NOT LD	31	81

n = 150

Total Agreement = 32 + 81 = 113

Total Disagreement = 6 + 31 = 37

Phi Coefficient = .27

Table 3

Class.LD2 and District 2 Placement Team Decisions

		PLACEMENT TEAMS: LD	PLACEMENT TEAMS: NOT LD
CLASS.LD2	LD	17	0
	NOT LD	15	22

N = 54

Total Agreement = 17 + 22 = 39

Total Disagreement = 0 + 15 = 15

Phi Coefficient = .56

Table 4

Class.LD2 and District 3 Placement Team Decisions

CLASS.LD2	PLACEMENT TEAMS: LD	PLACEMENT TEAMS: NOT LD
LD	15	4
NOT LD	22	19

N = 60

Total Agreement = 15 + 19 = 34

Total Disagreement = 4 + 22 = 26

Phi Coefficient = .24

Table 5

Class.LD2 and overall Expert Decisions in cases of Team/Class.LD2 Disagreement

		EXPERTS: LD	EXPERTS: NOT LD
CLASS.LD2	LD	22	8
	NOT LD	42	162

N = 78 Cases X 3 Experts = 234 Judgments

Total Agreement = 22 + 162 = 184

Total Disagreement = 8 + 42 = 50

Phi Coefficient = .46

Chi Square = 36.36

Table 6

Class.LD2 and Unanimous Expert Decisions in cases of Team/Class.LD2 Disagreement

CLASS.LD2 \ EXPERTS	LD	NOT LD
LD	5	0
NOT LD	6	47

n = 58

Total Agreement = 5 + 47 = 52

Total Disagreement = 0 + 6= 6

Phi Coefficient = .63

Chi Square = 23.38

Table 7

Correlation matrix: Phi coefficients describing the relationships between the judgements of teams, Class.LD2, and experts in cases where there is team/Class.LD2 disagreement

	CL	E 1	E 2	E 3
*PT	−1	−.39	−.59	−.24
**CL	—	.39	.59	.24
***E 1		—	.50	.79
E 2			—	.45
E 3				—

*Placement Teams

**Class. LD2

***Expert

N = 78

Table 8

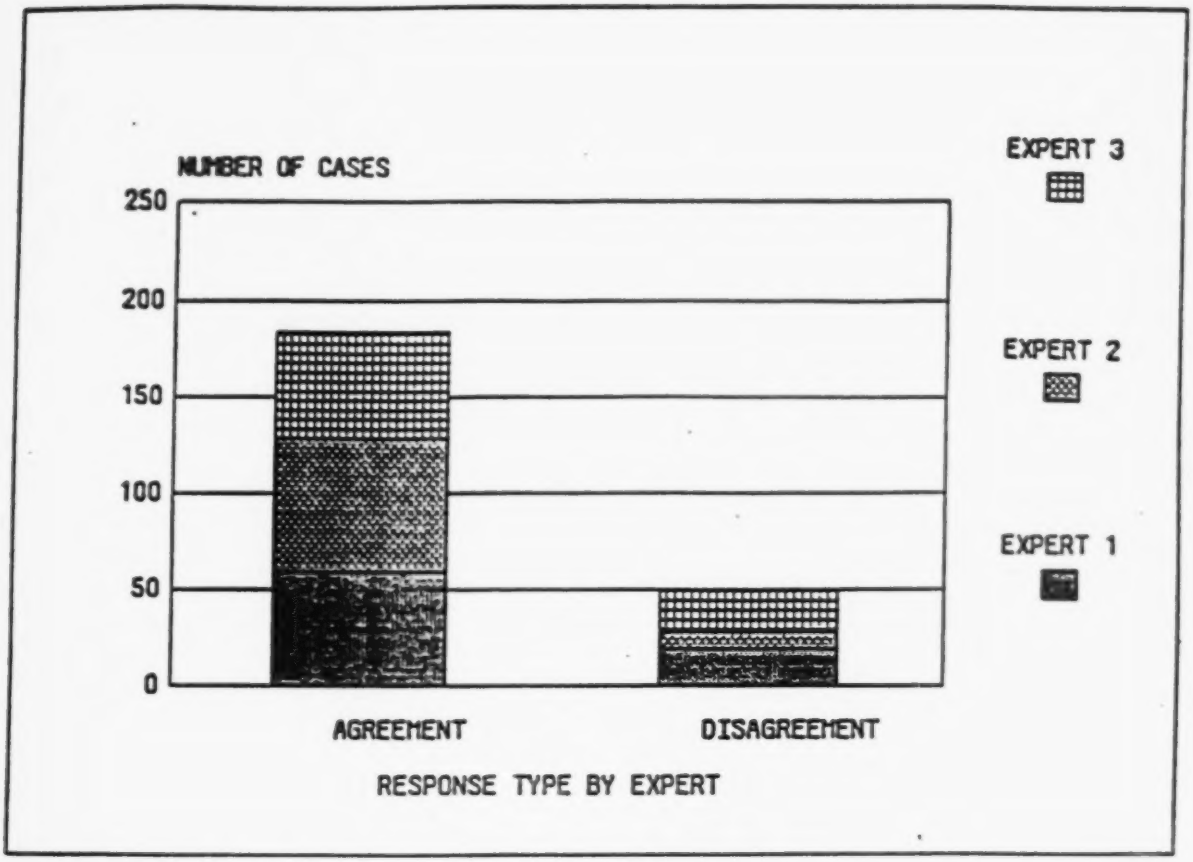
Intermediate Values of Critical Variables in Cases of Unanimous Expert/Class.LD2 Disagreement

Variable	IQ	Est. Discrepancy	Actual Discrepancy
Cutoff Point	85/84	40%	40%
Case # 1	100	49.3%	*39.6%
Case # 2	100	*39.5%	48.9%
Case # 3	*84	54.0%	55.8%
Case # 4	*83	50.0%	54.8%
Case # 5	104	40.0%	*38.5%
Case # 6	95	40.5%	*38.1%

* The critical score which failed to make Class.LD2's cutoff level.

Figure Caption

Figure 1. Number of agreements and disagreements of experts with Class.LD2. N = 78 X 3 = 234 judgments.


NUMBER OF CASES
250
200
150
100
50
0
AGREEMENT
DISAGREEMENT
EXPERT 3
EXPERT 2
EXPERT 1
RESPONSE TYPE BY EXPERT